Content-Based Image Retrieval in Medicine:

Retrospective Assessment, State of the Art, and Future Directions

L. Rodney Long,
National Library of Medicine, USA

Sameer Antani,
National Library of Medicine, USA

Thomas M. Deserno, and
Department of Medical Informatics, Germany

George R. Thoma
National Library of Medicine, USA

Abstract

Content-based image retrieval (CBIR) technology has been proposed to benefit not only the management of increasingly large image collections, but also to aid clinical care, biomedical research, and education. Based on a literature review, we conclude that there is widespread enthusiasm for CBIR in the engineering research community, but the application of this technology to solve practical medical problems is a goal yet to be realized. Furthermore, we highlight "gaps" between desired CBIR system functionality and what has been achieved to date, present for illustration a comparative analysis of four state-of-the-art CBIR implementations using the gap approach, and suggest that high-priority gaps to be overcome lie in CBIR interfaces and functionality that better serve the clinical and biomedical research communities.

Keywords

content-based image retrieval (CBIR); feature extraction; image indexing; medical image retrieval; performance; usability; Web user interface

INTRODUCTION

Content-based image retrieval (CBIR) technology exploits the visual content in image data. It has been proposed to benefit the management of increasingly large biomedical image collections as well as to aid clinical medicine, research, and education [1-2]. We treat CBIR as a set of methods that (1) index images based on the characteristics of their visual content, and (2) retrieve images by similarity to such characteristics, as expressed in queries submitted to the CBIR system. These characteristics, also referred to as "signature", may include intensity, color, texture, shape, size, location, or a combination of these. Sketching a cartoon, selecting an example image, or a combination of both methods, is typically used to form the query. The retrieved results are usually rank-ordered by some criteria; however, other methods, such as clustering of similar images, have been used to organize the results as well.

Practical application of CBIR depends on many different techniques and technologies applied at several stages in the indexing and retrieval workflow, such as: image segmentation and feature extraction; feature indexing and database methods; image similarity computation methods; pattern recognition and machine learning methods; image compression and networking for image storage and transmission; Internet technologies (such as JavaScript, PHP,

AJAX, Applet/Servlet); and human factors as well as usability. More recently, natural language processing has also been included, in attempts to exploit text descriptions of image content and the availability of standardized vocabularies [3]. It is through careful selection of appropriate methods from these fields that a successful CBIR application can be developed.

The technical literature regularly reports on experimental implementations of CBIR algorithms and prototype systems, yet the application of CBIR technology for either biomedical research or routine clinical use appears to be very limited. While there is widespread enthusiasm for CBIR in the engineering research community, the incorporation of this technology to solve practical medical problems is a goal yet to be realized. Possible obstacles to the use of CBIR in medicine include:

- The lack of productive collaborations between medical and engineering experts, which is strongly related to *usability* and *performance* characteristics of CBIR systems

- The lack of effective representation of medical *content* by low-level mathematical *features*

- The lack of thorough evaluation of CBIR system *performance* and its benefit in health care

- The absence of appropriate tools for medical experts to experiment with a CBIR application, which is again related to *usability* and *performance* attributes of CBIR systems.

Therefore, we take these four factors: content, features, performance, and usability as foundational in classifying and comparing CBIR systems, and in this paper we use these concepts as (1) an organizational principle for understanding the "gaps", or what is lacking in medical CBIR systems, (2) a lens for interpreting the main trends and themes in CBIR research over the past several years, and (3) a template for a systematic comparison of four existing biomedical CBIR systems.

The concept of gaps has often been used in CBIR literature, with the *semantic gap* being the most prominent example [1,2]. We have treated this "concept of gaps" as a paradigm for a broad understanding of what is lacking in CBIR systems and have extended the gap idea to apply to other aspects of CBIR systems [4], beyond the semantic gap. We may consider the semantic gap to be a *break or discontinuity* in the *aspect* of *image understanding*, with "human understanding" on one side of the gap and "machine understanding" on the other. Similarly, we may identify breaks or discontinuities in other aspects of CBIR systems, including the level of automation of feature extraction, with full automation on one side and completely manual extraction on the other; and, for another example, the degree to which the system helps the user refine and improve query results, with "intelligent" query refinement algorithms based on user identification of "good" and "bad" results on one side, and no refinement capability at all on the other. Each gap (1) corresponds to an aspect of a CBIR system that is explicitly or implicitly addressed during implementation; (2) divides that aspect between what is potentially a fuller or more powerful implementation from a less powerful one; and (3) has associated with it methods to bridge or reduce the gap.

MATERIALS AND METHODS

In order to assess medical CBIR retrospectively, we searched the web for relevant articles and identify the focus fields of past and current research. Using the concept of gaps [4], we also present the relevant differences in current medical CBIR systems illustratively, based on four state-of-the-art medical CBIR systems. Based on this analysis, we try to reliably predict future directions of medical CBIR, which we believe to be most important.

Retrospective Assessment

As a measure of types of research activity in the field of medical image retrieval, and of relative importance given to addressing particular system gaps, we surveyed the references to terms commonly used in the context of medical image retrieval in ten journals over the years 2001-2007. The journals were identified using informal selection criteria, but with the goal of providing a broad representation of the major publications reporting medical image retrieval research results. The journals and publishers are listed in Table 1. We followed a methodology similar to that discussed by Datta et al. [5], who carried out similar work for general image retrieval. Using Google Scholar (http://scholar.google.com), we searched for the terms {"medical image retrieval" AND *search_phrase*}, where *search_phrase* was one of the CBIR-related phrases given in Table 2.

State of the Art

In [4], we have identified a total of 14 gaps, and organized them into the basic "gap categories" given above: "Content Gaps," "Feature Gaps," "Performance Gaps," and "Usability Gaps." In addition to the gaps, other characteristics are useful to specify and distinguish medical CBIR systems. In [4], we group these under the general category of "system characteristics", which we further categorize as follows: (1) "intent and data" (the goal of using CBIR in the particular system, and the data that is used with it); (2) "input and output" (the specific I/O content); and (3) "feature and similarity" (the kind of features and distance measures used by the system).

The use of the concept of gaps, supplemented by system characteristics, has been proposed as a general methodology for comparative evaluation of CBIR systems, and for design planning in creating new systems. This conceptual organization is an effort towards encapsulating in a structured fashion the lessons learned in the published CBIR literature, and making system comparisons more comprehensive and practical.

In this paper, we illustrate the concrete application of these concepts to four state-of-the-art medical CBIR systems that are available online to the public via the Internet. All systems have been developed by at least one of the authors of this paper. This selection avoids problems that are generally associated with the judgment of work of other researchers.

A Viewpoint of Future Directions

Based on the retrospective literature review and comparative system overview, we suggest high-priority areas critical for moving CBIR into practical medical use. By its nature, this part is rather subjective and represents the personal viewpoints of the authors, rather than objective facts.

A RETROSPECTIVE ASSESSMENT OF MEDICAL CBIR SYSTEMS

The early years of medical CBIR have been reviewed by Müller et al [2]. As described in Section 2.1, we focus on the years 2001 through 2007. The number of citations returned for each of the search phrases is presented graphically in Figure 1.

What CBIR Researchers Have Emphasized

Inspection of Figure 1 shows, first of all, a high number of citations for the phrase "Content-Based Image Retrieval", which supports the idea that much of the medical image retrieval work in the engineering research community over the period investigated has in fact been related to CBIR. Other phrases near the high end of the citation scale suggest that most research attention has been in the areas of indexing, statistical methods, and learning methods. In terms of gaps addressed, the survey tends to support the view that most of the CBIR research effort over the

surveyed years has been in addressing the "Feature Gap Category", that is, the set of gaps dealing with the extraction of mathematical features from the images.

What CBIR Researchers Have Not Emphasized

At the lower end of the citation scale were the phrases referring to user interface, performance, interactivity, and relevance feedback. We note that while there were a relatively large number of references to "Web" in the journals, the considerably lower numbers of references to "user interface" suggest that many of the Web references did not refer to actual Web user interfaces, but more likely general acknowledgments of the significance of the Web. In terms of gaps not addressed, or weakly addressed, it appears that only a relatively small fraction of the CBIR research effort has been directed to addressing the "Performance Gap Category" and the "Usability Gap Category".

ILLUSTRATION: THE STATE OF THE ART OF MEDICAL CBIR SYSTEMS

In this section, we provide a concrete application of system analysis by gaps and system characteristics to four medical CBIR systems.

CervigramFinder

System Intent—The CervigramFinder system [6] operates on *cervicographic images* (also called *cervigrams*) and was created by the collaborative efforts of the National Cancer Institute (NCI) and the National Library of Medicine (NLM) for the study of uterine cervix cancer. This cancer is closely related to the chronic infection of certain types of Human Papillomavirus (HPV). To visually screen for pre-invasive cervical lesions or for cancer, one cost-effective method is cervicography. Cervicographic screening is based on the acetowhitening phenomenon: HPV-infected abnormal tissue often turns white after being treated with 3-5% acetic acid. A cervigram is a 35-mm photograph of the cervix taken approximately one minute after acetic acid exposure. NLM has created a cervigram database containing approximately 100,000 cervigrams taken during two major projects in cervical cancer carried out by NCI to study the natural history of HPV infection and cervical neoplasia, the *Guanacaste* and *ALTS* projects [7-8]. In addition to cervigrams, correlated clinical, cytologic and molecular information were also collected.

Interface—CervigramFinder operates on a subset of the cervigram database. To use this system, the user defines a query region by marking a region of interest on an image through the graphical user interface shown in Figures 2a and 2b. (In the query shown in these Figures, the user is searching on the "location" feature and is limiting the search to regions that already have the semantic labeling "AW", for "acetowhitened".) The system then (1) calculates the feature vector of the query region for the specified features and (2) compares that query feature vector with the pre-computed feature vectors of regions stored in the database. The returned regions, shown on their parent images, are ranked by the degree of their similarity to the query feature vector and presented on a multi-image display, along with associated text information. The (visual) features used are color, texture and size. Shape is significantly less important as a feature for identifying or distinguishing regions in this application since these tissue types do not exhibit any particular shape except for the *os* regions (the os is the opening into the uterus) which are somewhat elliptic. In order to facilitate system evaluation by medical experts located at geographically different sites, as well as to allow the final system to be accessed remotely for either diagnosis or education in the future, the system is implemented using a distributed client/server framework.

Gaps and System Characteristics—Gap and system characteristics of CervigramFinder are given in Tables 3 and 4, respectively, which provide a side-by-side comparison with

corresponding gaps and characteristics of the CervigramFinder, SPIRS, IRMA, and SPIRS-IRMA systems. Significant gaps that are yet to be addressed in the CervigramFinder system include, for Feature Gaps, lack of multiscale analysis (only single-scale is used) ; for Performance Gaps, lack of integration into use in a biomedical system and lack of database indexing; for Usability Gaps, neither user feedback on relative similarity of returned images, and nor query refinement is provided. Capabilities that have at least partially addressed some gaps include, for Content Gaps, semantic labeling of regions in the database images; for Feature Gaps, some computer-assisted feature extraction (for indexing features, a user must manually mark boundaries of significant regions; algorithms then compute mathematical features from these regions); for Performance Gaps, online implementation and qualitative retrieval evaluation; and, for Usability Gaps, retrieval by both user selection of pre-stored regions-of-interest ("query by composition") and by interactive user sketch. We also note that CervigramFinder has been exercised by several medical experts with their system interactions digitally recorded, for improvement of usability of the system. The system characteristics of CervigramFinder indicate that it is for research, teaching, and learning; it uses 2D data; it operates only on image data, both for input and output. We note also that CervigramFinder operates on color image data, making it unique in that respect among the four systems that we discuss.

SPIRS

System Intent—The Spine Pathology & Image Retrieval System (SPIRS) [9] was developed at the U. S. National Library of Medicine to retrieve x-ray images from a large dataset of 17,000 digitized radiographs of the spine and associated text records. Users can search these images by providing a sketch of the vertebral outline or selecting an example vertebral image and some relevant text parameters. Pertinent pathology on the image/sketch can be annotated and weighted to indicate importance. This hybrid text-image query yields images containing similar vertebrae along with relevant fields from associated text records, which allows users to examine the pathologies of vertebral abnormalities.

Interface—SPIRS provides a Web-based interface for image retrieval using the morphological shape of the vertebral body. A query editor enables users to pose queries either by sketching a unique shape, or by selecting or modifying an existing shape from the database. Additional text fields enable users to supplement visual queries with other relevant data (e.g., anthropometric data, quantitative imaging parameters, patient demographics). These hybrid text-image queries may be annotated with pertinent pathologies by selecting and weighting local features to indicate importance. Query results appear in a customizable window that displays the top matching results and related patient data. The SPIRS interface is shown in Figure 3.

Gaps and System Characteristics—Significant gaps that are yet to be addressed in the SPIRS system are similar to those for CervigramFinder, and include, for Feature Gaps, lack of multiscale analysis; for Performance Gaps, lack of integration into use in a biomedical system and lack of quantitative evaluation; for Usability Gaps, no user query refinement. (However, see comments about "data exploration" below.) Capabilities that have at least partially addressed some gaps include, for Content Gaps, manual labeling of vertebrae by anatomical type; for Feature Gaps, computer-assisted feature extraction (an Active Contours algorithm is used to find approximate boundaries of vertebrae in the images; these boundaries then are manually reviewed and corrected); for Performance Gaps, feature vector indexing by K-D Tree, and qualitative evaluation; and, for Usability Gaps, support for both query by composition (see Section 2.1.1) and by interactive user sketch. We also note that SPIRS provides capability to specify not only the shape to be used in the query, but *which part of the shape should be used*, so that the user may focus on the fine level of structure that is often critical in biomedical

image interpretation. In addition, SPIRS provides (1) "basic" user feedback on each returned image, namely, a figure of dissimilarity to the query image; and (2) a "data exploration" capability, which takes query results as a beginning point to initiate new and related queries; using a given query result, that is, a vertebral shape returned by a query, the entire spine containing that shape may be displayed; then the user may select a vertebra in that same spine and use its shape as a new query. It should be noted that SPIRS, like CervigramFinder, *operates on local, region-of-interest data* in the image. The system characteristics of SPIRS indicate that it is for research, teaching, and learning on 2D data; it accepts as input, and creates output "hybrid" data (both text and image). In this regard, SPIRS allows the user to specify as a query a vertebral shape and some text (such as age, race, gender, presence/absence of back or neck pain, and vertebra tags such as "C5", to indicate the class of vertebrae being searched for). It then returns such text, along with the associated image data.

IRMA

System Intent—The Image Retrieval in Medical Applications (IRMA) *project* [10-12] has the following goals: (1) automated classification of radiographs based on global features with respect to imaging modality, body orientation with respect to the x-ray beam (e.g., "anterior-posterior" or "sagittal"), anatomical body region examined, and the biological system under investigation; and (2) identification of local image features including their constellation within a scene, which are relevant for medical diagnosis. These local features are derived from a priori classified and registered images that have been segmented automatically into a multi-scale approach. IRMA analyzes content of medical images using a six-layer information model: (1) raw data, (2) registered data, (3) feature, (4) scheme, (5) object, and (6) knowledge.

The IRMA *system* that is currently available via the Internet retrieves images similar to a query image with respect to a selected set of features. These features can, for example, be based on the visual similarity of certain image structures. Currently, the image data consists of radiographs. It uses a reference database of 10,000 images categorized by image modality, orientation, body region, and biological system.

Interface—The system architecture has three main components: (1) the central database, containing images, processing schemes, features, and administrative information about the IRMA workstation cluster; (2) the scheduler, which balances the computational workload across the cluster; and 3) the Web server, which provides the graphical user interface to the IRMA system for data entry and retrieval. Extended query refinement is established by logging all user interaction in the system database that also hold the features extracted from the images [12]. The IRMA system interface is shown in Figure 4.

Gaps and System Characteristics—In contrast to the rather general concept within the IRMA project, the IRMA system that is currently demonstrated on the web has some significant gaps that are still yet to be addressed. These gaps include, for Content Gaps, lack of semantic labeling; for Feature Gaps, only operation on global image characteristics is supported, and multiscale analysis is lacking; for Performance Gaps, lack of integration into use in a biomedical system, lack of feature vector indexing, and lack of quantitative evaluation. Capabilities that at least partially address system gaps include, for Feature Gaps, fully automatic feature extraction (facilitated, of course, by the fact that IRMA operates on the image as a whole, so that segmentation of particular regions-of-interest prior to feature extraction is not required); for Performance Gaps, a widely-publicized and mature online Internet presence, and qualitative retrieval evaluation; and, for Usability Gaps, an extremely flexible query refinement mechanism that lets the user step back and forth among queries done in a session, and lets the user combine queries with union, intersection, and negation operators. This is coupled with an advanced feedback measure that assists the user in judging how closely a

retrieved image matches not only a single image used in the query, but how closely it matches a weighted set of images. The system characteristics of IRMA indicate that it is for research, teaching, and learning use on 2D data.

SPIRS-IRMA

System Intent—IRMA, described above, aims at providing visually rich image management through CBIR techniques applied to medical images using intensity distribution and texture measures taken globally over the entire image. This approach permits queries on a heterogeneous image collection and helps identify images that are similar with respect to global features, e.g., all chest x-rays in the AP (anterior-posterior) view. However, the IRMA system lacks the ability to find particular pathology that may be localized in specific regions within the image. In contrast, the SPIRS system provides localized vertebral shape-based CBIR methods for pathologically sensitive retrieval of digitized spine x-rays and associated metadata. In the SPIRS system, the images in the collection must be homogeneous, i.e., a single modality imaging the same anatomy in the same view, e.g., vertebral pathology expressed in spine x-ray images in the sagittal plane. Observing the different strengths of the two systems led to the idea of combining these complementary technologies to create an SPIRS-IRMA system [13] that will eventually support both whole image and local feature-based retrieval so that users may find images that are not only similar in overall appearance but also similar with respect to locally-expressed pathology.

Interface—Initial work toward creating such a system has begun and some capabilities are in place; the current SPIRS-IRMA interface is shown in Figure 5.

Gaps and System Characteristics—SPIRS-IRMA, then, is an example of combining the capabilities of different CBIR implementations, developed by different research groups, as a strategy of closing CBIR gaps of the individual systems. We noted above that the IRMA system operates on global image data only, while the SPIRS system operates only on local region-of-interest image data that has been segmented from the image. The SPIRS-IRMA system is the first step toward a system that will integrate the capabilities of these two systems. At the current time, the SPIRS capabilities for retrieval by vertebrae shape similarity, and the SPIRS vertebrae shape database, have been coupled to the IRMA user interface, so that an IRMA user has full access to SPIRS for vertebrae retrieval by shape. A user may log in to the IRMA system and access an interface that enables the retrieval of spine vertebrae by shape. This capability uses the combined resources of servers operating in Germany (Aachen) and the U.S. (Bethesda, Maryland) which are linked through an XML-based service protocol that is used to coordinate the transmission of the query and the query results between the servers.

This system lays the groundwork to perform global image searches to identify images of interest, and then to use local region-of-interest search capability to drill down into specific localized anatomy or pathology. It already combines the IRMA interface (with session query management), with the local region search capability of the SPIRS system.

While the goal is for the SPIRS-IRMA system to eventually possess all of the strengths of both systems, the current, initial system, provides only some of these capabilities. Also, some of the individual system strengths are not available in the current SPIRS-IRMA implementation (for example, SPIRS returns both images and keywords, but SPIRS-IRMA returns only images). Significant gaps yet to be addressed in the SPIRS-IRMA system include the following: no semantic content is available to the user (the manual semantic labeling of SPIRS is not yet available under SPIRS-IRMA), the image structure that may be used in queries is only local, as in SPIRS, at the current time, and only query by composition (pre-stored shapes) is available (SPIRS-IRMA does not allow interactive sketch). A gain over the SPIRS system, though, has

been the narrowing of Performance and Usability gaps through the use of the well-known IRMA interface, and by the versatility of its session management capabilities available for searching the SPIRS data. The future joining of the two systems to create image search by both global and local characteristics will add capability that is rare if not unique in the medical CBIR field. The system characteristics of SPIRS-IRMA indicate that its use is for research, teaching, and learning use on 2D data.

DISCUSSION: FUTURE DIRECTIONS FOR MEDICAL CBIR

Creating effective collaborations among different, geographically-separated CBIR engineering research groups, and collaborations among the engineering and medical communities to advance this field, will likely remain a challenge for the foreseeable future. Nevertheless, certain efforts within the engineering community are worth noting, including (1) the important Image-CLEF competition [3], which allows evaluation of algorithmic approaches of multiple research groups on a single image test set, (2) the convening of CBIR workshops at professional conferences, such as those held at MICCAI in 2007 [14] and SPIE Medical Imaging in 2008 [15], (3) the collection of segmentation data from medical experts, (4) the exposure of CBIR systems to medical experts, though in small scale efforts to date, and (5) collaborative work to combine and make different CBIR systems interact, typified by SPIRS-IRMA, to exploit the strengths of the individual systems.

Effectively representing medical content by low-level mathematical features is essentially grappling with the semantic gap, which may possibly remain a perennial problem. This does not mean, however, that tools for retrieval by image content may not be made increasingly effective. Easy-to-use relevance feedback mechanisms, such as those supported by the IRMA system, ameliorate this situation somewhat by allowing the user to quickly refine queries by identifying specific returned results as desirable or not desirable. Our literature search suggests that this entire domain of relevance feedback has been under-researched, and we anticipate considerable room for growth and improvement of existing techniques.

Evaluation of CBIR systems has been a particularly difficult issue, with precision and recall measures frequently being used, but with a "ground truth" which may reflect a high degree of variability in expert opinion. The crucial threshold for medical CBIR system evaluation remains, of course, not a quantitative mark defined in the engineering environment, but the degree of usefulness to the biomedical community in such systems becoming truly valuable aids in clinical and research problem-solving.

It is common for engineering groups engaged in CBIR development to express a desire for closer collaboration with the medical community. It is less common to propose solutions for bridging this *collaboration gap*. We suggest more proactive steps to expose CBIR tools to the medical community as an effort to help overcome this problem. This entails both (1) understanding the types of biomedical problems for which CBIR can potentially have a clinical or research impact, and (2) tailoring tool interfaces to operate in the "patient-centric" mode of the medical environment; with the appropriate balance of simplicity and power, as judged by the medical user; with labeling and terminology appropriate for the medical user; and with interface capabilities for importing and exporting information from other data sources that are important to the medical user.

CONCLUSION

Success of a particular technology is often due to the confluence of available, supporting technologies at the time of critical need. Content-Based Image Retrieval of medical images has achieved a degree of maturity, albeit at a research level, at a time of significant need.

However, the field has yet to make noticeable inroads into mainstream clinical practice, medical research, or training. In this article we have explored the field through the concept of gaps or shortcomings in comparison with an idealized system. By addressing and minimizing these gaps, a system may be better positioned for use in the biomedical world. We have characterized CBIR system gaps under the broad categories of content, feature, performance, and usability and suggest that the published CBIR technical literature reflects too little attention to closing the gaps of performance and usability, although these are perhaps the gating factors that limit closer collaboration with the biomedical community. We suggest early, proactive system design incorporating the workflow, terminology, and modes of operation of the biomedical user as a needed effort for enhancing collaboration with the medical community.

Acknowledgments

REFERENCES

1. Antani, S.; Long, LR.; Thoma, G. Bridging the Gap: Enabling CBIR in Medical Applications. Proc. 21st International Symposium on Computer-Based Medical Systems (CBMS 2008); University of Jyväskylä, Finland. Jun. 2008 (Accepted)

2. Müller H, Michoux N, Bandon D, Geissbuhler A. A Review of Content-Based Image Retrieval Systems in Medical Applications - Clinical Benefits and Future directions. Int J Med Inform Feb;2004 73(1): 1–23. [PubMed: 15036075]

3. CLEF-Campaign. http://www.clef-campaign.org

4. Deserno, TM.; Antani, S.; Long, R. Ontology of Gaps in Content-Based Image Retrieval. Journal of Digital Imaging. Feb. 2008 *Online*: http://www.springerlink.com/content/?k=deserno

5. Datta R, Joshi D, Li J, Wang JZ. Image Retrieval: Ideas, Influences, and Trends of the New Age. ACM Transactions on Computing Surveys April;2008 Vol. 40(No. 2):1–66.

6. Xue, Z.; Long, LR.; Antani, S.; Jeronimo, J.; Thoma, GR. Proc. SPIE Medical Imaging 2008. Vol. vol. 6919. San Diego, CA: Feb. 2008 A Web-accessible content-based cervicographic image retrieval system; p. 691907-1-9

7. Herrero R, Schiffman MH, et al. Design and methods of a population-based natural history study of cervical neoplasia in a rural province of Costa Rica: the Guanacaste Project. Pan American Journal of Public Health 1997;1(15):362–375. [PubMed: 9180057]

8. Schiffman M, Adrianza ME. ASCUS-LSIL Triage Study. Design, methods and characteristics of trial participants. Acta Cytologica Sep-Oct;2000 44(5):726–42. [PubMed: 11015972]

9. Hsu W, Antani SK, Long LR. SPIRS: A Framework for Content-based Image Retrieval from Large Biomedical Databases. Proc. of MEDINFO August;2007 12(1):188:92.

10. IRMA Project. http://www.irma-project.org

11. Lehmann TM, Guld MO, Thies C, Fischer B, Spitzer K, Keysers D, Ney H, Kohnen M, Schubert H, Wein BB. Content-based image retrieval in medical applications. Methods of Informatics in Medicine 2004;43:354–61.

12. Deserno TM, Güld MO, Plodowski B, Spitzer K, Wein BB, Schubert H, Ney H, Seidl T. Extended query refinement for medical image retrieval. Journal of Digital Imaging. 2007 online-first, DOI 10.1007/s10278-007-9037-4.

13. Antani SK, Deserno TM, Long LR, Güld MO, Neve L, Thoma GR. Interfacing global and local CBIR systems for medical image retrieval. Proc. Workshop on Medical Imaging Research (Bildverarbeitung fur die Medizin) March;2007 :166–71.

14. Content-based Image Retrieval for Biomedical Image Archives: Achievements, Problems, and Prospects, MICCAI 2007. Oct. 2007 http://www.eng.tau.ac.il/~hayit/MICCAI_CBIR_workshop/

15. Content-Based Image Retrieval: Major Challenges for Medical Applications. SPIE Medical Imaging 2008. Feb. 2008 Presentations at: http://archive.nlm.nih.gov/spiemi08/CBIRpanel.php
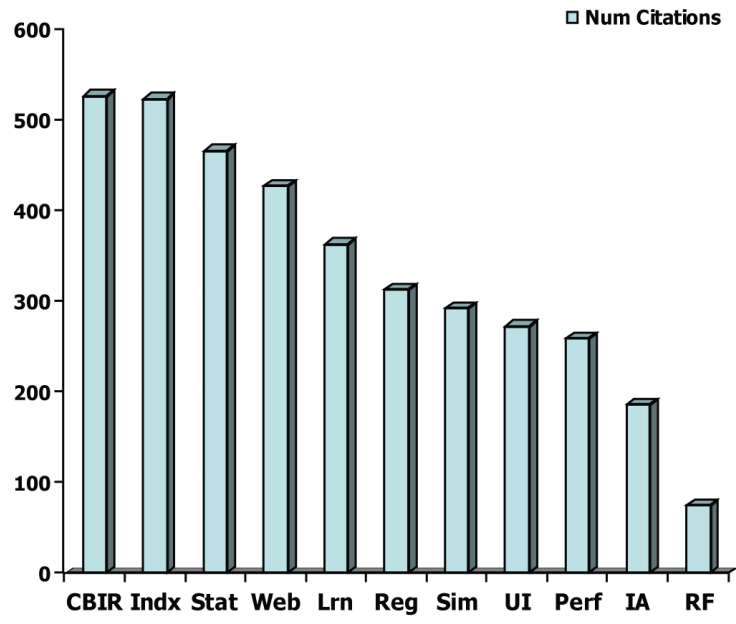
Figure 1.
Journal citation results (for phrases related to medical image retrieval) for journals surveyed 2001-2007. For list of journals, see Table 1. For explanation of abbreviations on x-axis, see Table 2.
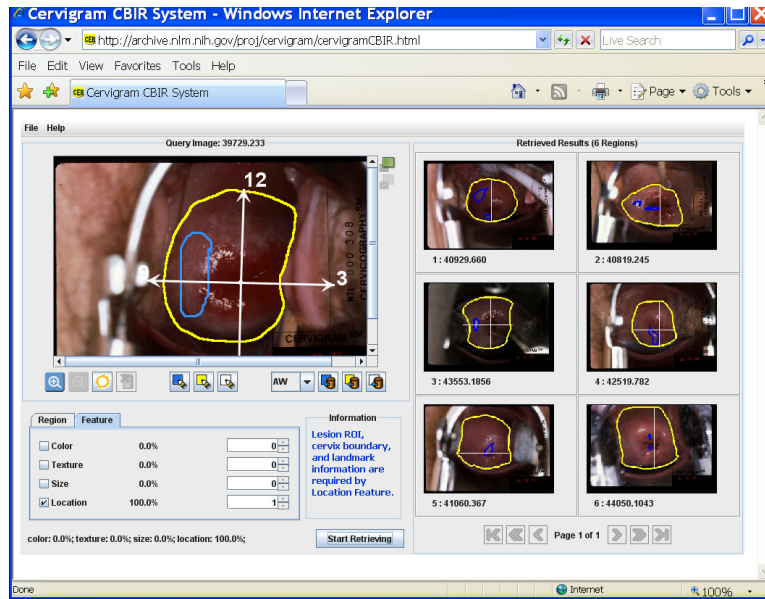
Figure 2a.
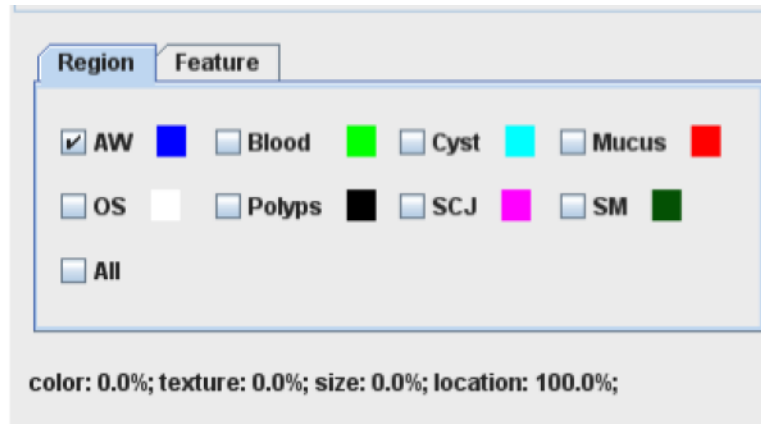CervigramFinder interface; "feature" panel in lower left shows that user is searching on "location"

Figure 2b.
CervigramFinder interface "region" panel, showing that user may limit search to semantically-labeled region types
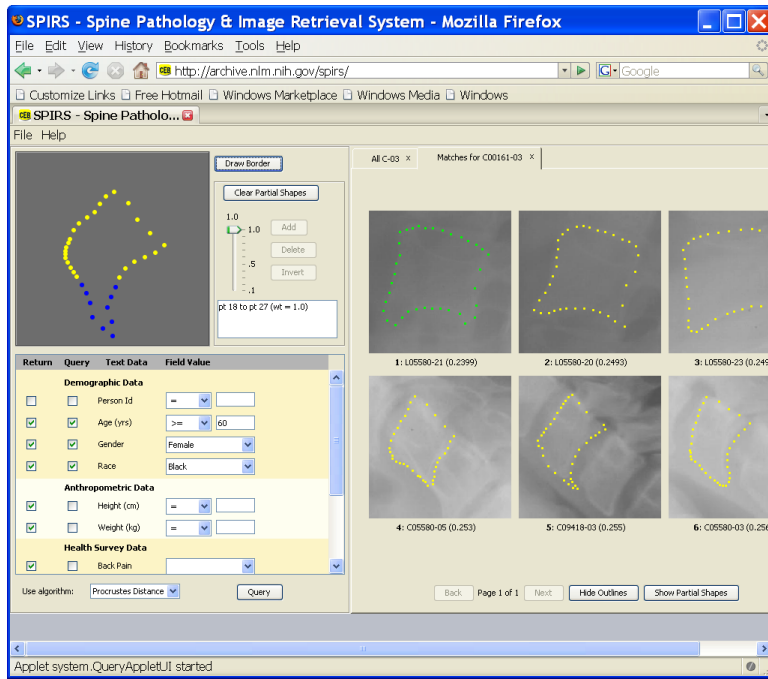
Figure 3.
SPIRS interface; example query for records satisfying criteria {(age >=60, gender=female, race=black) AND having vertebrae similar to lower/front of sketch)
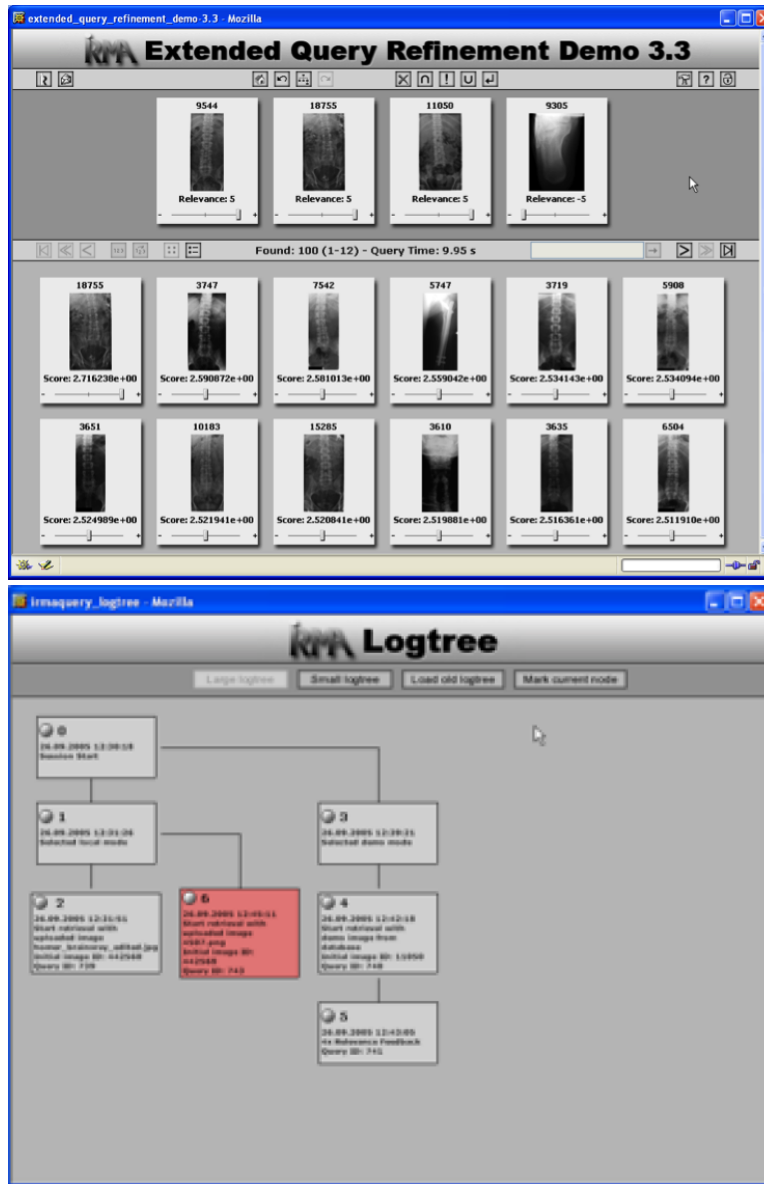
Figure 4.
a. IRMA query interface with relevance feedback. The initial query image was user-uploaded from the user's computer.
b. The IRMA session logging provides complete access to previous session states.
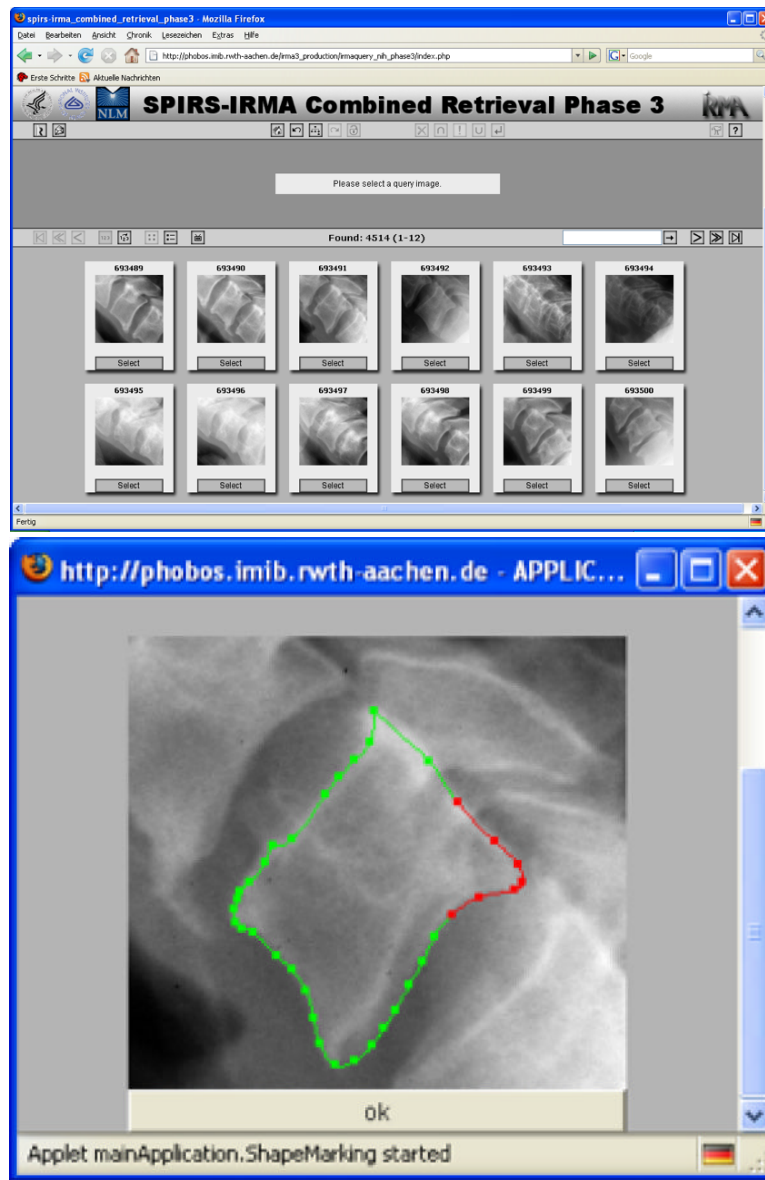
Figure 5.
a. SPIRS-IRMA interface for searching vertebra shapes
b. A vertebra shape is represented by 36 landmark points, and the user can select a partial shape of interest

Table 1

Journals surveyed for medical image retrieval terms

	Journal	Publisher
1	*Computerized Medical Imaging and Graphics*	Elsevier Science, Amsterdam, The Netherlands
2	*IEEE Transactions on Image Processing*	The Institute of Electrical and Electronics Engineers (IEEE); IEEE Press, Piscataway, NJ, USA
3	*IEEE Transactions on Knowledge and Data Engineering*	IEEE Press, Piscataway, NJ, USA
4	*IEEE Transactions on Medical Imaging*	IEEE Press, Piscataway, NJ, USA
5	*IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*	IEEE Press, Piscataway, NJ, USA
6	*International Journal of Medical Informatics*	Elsevier Science, Amsterdam, The Netherlands
7	*Journal of the American Medical Informatics Association (JAMIA)*	Hanley & Belfus, Inc., Orlando, FL, USA
8	*Journal of Digital Imaging*	Springer, New York, NY, USA
9	*Journal of Electronic Imaging*	Society of Photooptical Instrumentation Enginering (SPIE); SPIE Press, Bellingham, WA, USA
10	*Radiology*	Radiological Society of North America (RSNA), Oak Brook, IL, USA

Table 2

Search phrases ANDed with "medical image retrieval", in decreasing order of number of citations. Abbreviations are used in Figure 1.

	Search Phrase	Abbreviation
1	Content-Based Image Retrieval	CBIR
2	Indexing	Indx
3	Statistical	Stat
4	Web	Web
5	Learning	Lrn
6	Registration	Reg
7	Similarity	Sim
8	User Interface	UI
9	Performance	Perf
10	Interactivity	IA
11	Relevance Feedback	RF

Table 3

System Gaps Compared Across CBIR Systems. (See reference [4] for complete explanation of terms. A few of the most significant are given here.) Semantic/ Manual: some semantic labeling of image content; Use Context/Narrow: one or small number of image modalities; Scale/Single: no multiscale processing; Query/Composition: pre-stored shapes or patterns are used; Query/Sketch: interactive user sketch; Feedback/Basic: only similarity or dissimilarity to single query image is provided; Feedback/Advanced: measure of match to weighted image set is provided; Refinement/Complete Combination: complete query history in session is maintained and queries may be combined

		CervigramFinder	SPIRS	IRMA	SPIRS-IRMA
Content	Semantic	Manual	Manual	Not addressed	Not addressed
	Use Context	Narrow	Narrow	Narrow	Narrow
Feature	Extraction	Computer-assisted	Computer-assisted	Automatic	Computer-assisted
	Structure	Local	Local	Global	Local
	Scale	Single	Single	Single	Single
	Space+Time Dimension	Not applicable	Not applicable	Not applicable	Not applicable
	Channel Dimension	Not applicable	Not applicable	Not applicable	Not applicable
	Application	Online	Online	Online	Online
	Integration	Not addressed	Not addressed	Not addressed	Not addressed
Performance	Indexing	Not addressed	Software supported (K-D Tree)	Not addressed	Software supported (K-D Tree)
	Evaluation	Qualitative—900 cervigrams	Qualitative—4,514 vertebrae x-rays	Qualitative—10,000 radiographs	Qualitative—4,514 vertebrae x-rays
	Query	Hybrid (Composition, Sketch)	Hybrid (Composition, Sketch)	Pattern	Composition
Usability	Feedback	Not addressed	Basic	Advanced	Advanced
	Refinement	Not addressed	Not addressed	Complete Combination	Complete Combination

Table 4

Comparative system characteristics

		CervigramFinder	SPIRS	IRMA	SPIRS-IRMA
Intent & Data	System Intent	Hybrid (Research, Teaching, Learning)	Hybrid (Research, Teaching, Learning)	Hybrid (Research, Teaching, Learning)	Hybrid (Research, Teaching, Learning)
	Data Domain	2D	2D	2D	2D
	Data Range	2D	2D	2D	2D
Input & Output	Input Data	Image	Hybrid (Image, Keyword)	Image	Image
	Output Data	Image only	Hybrid (Image, Keyword)	Image only	Image only
Feature & Similarity	Image Features	Hybrid (Color, Texture, Special -Location)	Grayscale	Grayscale	Grayscale
	Distance Measure	Metric - Euclidean	Metric - Euclidean	Metric - Euclidean	Metric - Euclidean